DUPLICATE ALSO

0290

This document is unpublished.
It must not be quoted in print
without permission from the
appropriate Met Office Branch.

CENTRAL FORECASTING TECHNICAL NOTE NO.9

THE EFFECTIVENESS OF GRIB PACKING FOR WAFS

BY

G.H.ROSS. P.J.SMITH AND S.D.TURNER

OCTOBER 1992

METEOROLOGICAL OFFICE.

Central Forecasting Technical Note No.9

The effectiveness of GRIB packing for WAFS.

Bracknell::::1992:Pp.i+4+App.5:30cm:

92110861 551.509.313.21

Headquarters, Bracknell

FH3A 19 NOV 1992
National Meteorological Library
London Rd, Bracknell, RG12 2SZ

METEOROLOGICAL OFFICE
160254
C
LIBRARY

CENTRAL FORECASTING TECHNICAL NOTE NO.9

THE EFFECTIVENESS OF GRIB PACKING FOR WAFS

BY

G.H.ROSS, P.J.SMITH AND S.D.TURNER

CENTRAL FORECASTING DIVISION

U.K. MET OFFICE

OCTOBER 1992

Central Forecasting Division
Meteorological Office
London Road
Bracknell Berkshire
RG12 2SZ

N.B. This paper has not been published. Permission to quote form it should be obtained from the Assistant Director of the above Meteorological Office Division.

The Effectiveness of GRIB Packing for WAFS.

Central Forecasting Division
U.K. Met Office

This report considers the size of a forecast data file packed by the WMO GRIB code FM92-IX (known as GRIB Version 1) when different resolutions and packing strategies are used.

Current forecast products are on a 2.5 degrees latitude by 5.0 degrees longitude. Proposals for future resolutions are 1.25x2.5, 1.25x1.25, and an effective resolution of 1.25x1.25 at the equator, gradually thinning until there are only 4 points at the pole. This last grid is known as the "Thinned" grid, and is packed as an "Irregular" grid in the GRIB V1 code.

The results show that the "Thinned" grid can meet the criterion of around 3 Mbytes per Forecast which is the approximate size of the 1.25x2.5 grid packed under a previous version of the GRIB code, known as Version 0.

Thinned Grid.

Figure 1 shows the thinned grid plotted on a latitude-longitude projection. (The lat-lon projection is useful only for such display purposes). The grid has a decreasing number of points per latitude row as it nears the pole, with four points at each pole. There are always points at 0E, 90E, 180E and 270E, so that the grid could be split into 4 sections per hemisphere, each of which has the same point pattern and a point at the pole. (The original specification had a single pole point at each pole.)
This study compares GRIB version 1 packing sizes for a regular 1.25x1.25 grid and a regular 1.25x2.5 grid using 3 packing methods, against the thinned grid with first order and second order packing of an irregular grid.

Forecast Data.

All the fields packed are Whole Globe fields, with forecast data taken from a particular day and forecast lead time in the Northern Hemisphere summer. The size of field varies with season, fluctuating by less than 10%, and is largest at this time of the year. The annual variation is due to the stronger Southern Hemisphere winter jets.

GRIB packing has header information which varies between 83 bytes for simple packing, regular grid, and 373 for simple packing, thinned grid with a grid descriptor. Splitting the globe into sectors will only expand the size by a small amount, so although there is one header per field in the following results, splitting makes no difference to our conclusions.

Fields at 250mb and 850mb were considered to give examples of different scales at different levels. Temperature fields are larger at 850 mb, but wind fields are smaller. The parameters analysed are Height, Temperature, and wind split into U and V wind components. Normally height fields are not sent to

airlines in the WAFS code.
Wind Speed and Direction were considered in parallel and are not presented here. Although wind speed packs well, direction does not, and the result is very similar to U and V. When we consider the effects of precision errors on cross winds calculated using data in both representations, U and V values have largest errors at low wind speeds, while with speed and direction it occurs at high wind speeds. Although a precision can be chosen to minimise any error, U and V will be used for WAFS.

Tables.

Table 1 a) and b) are for the packing of the 1.25x2.5 grid, 2 a) and b) for the 1.25x1.25 regular grid, and 3 a) and b) are for the thinned grid.
Table 4 is a summary of the size of a whole forecast for the best choice of parameter precision.
The thinned grid has 26616 points (Tables 3 a) and b) are for the first specification with a single pole point), the regular 1.25x1.25 grid has 41760 points and the 1.25x2.5 grid has 20880 points.
Actually the GRIB code assumes only one point at the pole, so the true number of points packed for the 1.25x1.25 regular grid, for example, is not 41760 but 41186 points. We consider true packing sizes.

The set of a) tables are results for whole globe fields at 850mb, and the set b) are for 250mb.

Within each table the sub-tables for the parameters, Height, Temperature, and the U and V components of the wind, packed separately. In each sub-table the size in bytes is listed for different parameter precisions in columns of packing method.

Packing Methods.

Within each parameter there are a range of packing precisions. The values are in powers of 2 in the units of the parameter. GRIB allows fixed precision, where each point is allocated a fixed number of bits (and therefore the precision varies from field to field) or fixed precision (and the bit size varies). These are "first order packing" which we call "simple packing". Grib release 1 allows a "second order packing". We have chosen "row by row" packing, where each line of latitude is packed separately to a fixed precision.

For the regular grid, we have used "simple packing", "mapped repetition", and "row by row" packing. The other method is "simple packing, fixed bit length", but that gives one of the values in the simple packing list. For the Thinned Grid we considered "simple" and "row by row". These are described below:

Simple Packing.

A packing precision is chosen as a power of 2 on the unit value. For example, the height field has been sampled in discrete steps of 1 dkm, 0.5, 0.25 and 0.125 dkm.
The whole field is scanned to find the smallest and the largest values. This gives the bit size necessary for the packing because:

No. of bits = (largest-smallest)/precision.
....rounded up.

Each value in the field is stored as the number of increments larger than the lowest value.

Simple Packing - Fixed Bit Length.

This is simple packing, but the number of bits is preset, and the consequent packing precision derived from the relationship above.
The data sent from the UK using GRIB version 0 has this packing, with a choice of 12 bits for height fields and pressure values, and 8 bit precision for the rest.

Mapped Repetition.

This is not a standard packing technique, but has been adopted to give some idea what extra packing that run length encoding might contribute, because no run length encoding algorithm is currently available for GRIB code.
After a field has been "simple packed" it is scanned to find runs of the same value. Second and further values are treated as if they were missing, and a bit map of these missing values is stored in the GRIB version 1 header information. The assumption is that a decoding routine replaces a value which is "missing" with the previous value. So, effectively, a repeated value is stored with a single bit.
Obviously this only has advantages where the data is slowly varying compared to the packing precision.

Row by Row Packing.

The packing again is simple packing. The difference is that each latitude row is packed separately. The advantage is gained because on very few rows will there be as large a range as from the lowest to the highest values over the whole field. For example the temperature field varies strongly with latitude, but only infrequently is there as large a range *around* a latitude. Over the pole, where there is a considerable amount of redundant information, there will seldom be a strong temperature gradient. This means that the excessive number of points probably can be packed with only a few bits per point.

Results.

For temperature, which is a slowly varying parameter, the mapped repetition scores very well. It works even better when the precision is low. However slow variation also gives a small range over a latitude row, so row-by-row packing works well too. In fact for the regular grid, row-by-row packing is best overall, and is only occasionally beaten by mapped repetition.

Comparing results on the 1.25x2.5 grid and the 1.25x1.25 grid, although for simple packing, the difference is (of course) close to a factor 2, for row-by-row packing there is a slight extra packing for the higher resolution, and the ratio is about 1:1.9 .

When we compare 1.25x2.5 resolution using simple packing and thinned grid using row by row, then the size is very similar. Temperature fields are smaller for the thinned grid and wind components larger at 250mb, and this reverses for 850 mb. Heights are always a little smaller for the lower resolution.

We have some freedom to specify the precision of the data sent to the airlines. Because there may be other users who would like a high precision, as high a precision as possible should be considered. Height at 0.5dkm, Temperature at 0.25C, Wind components at 0.25m/s are good choices.

Table 4 summarises these choices in terms of the full file size for each forecast. We assume that there are 11 levels in the vertical, 4 forecast times, and 3 parameters (T, U and V).

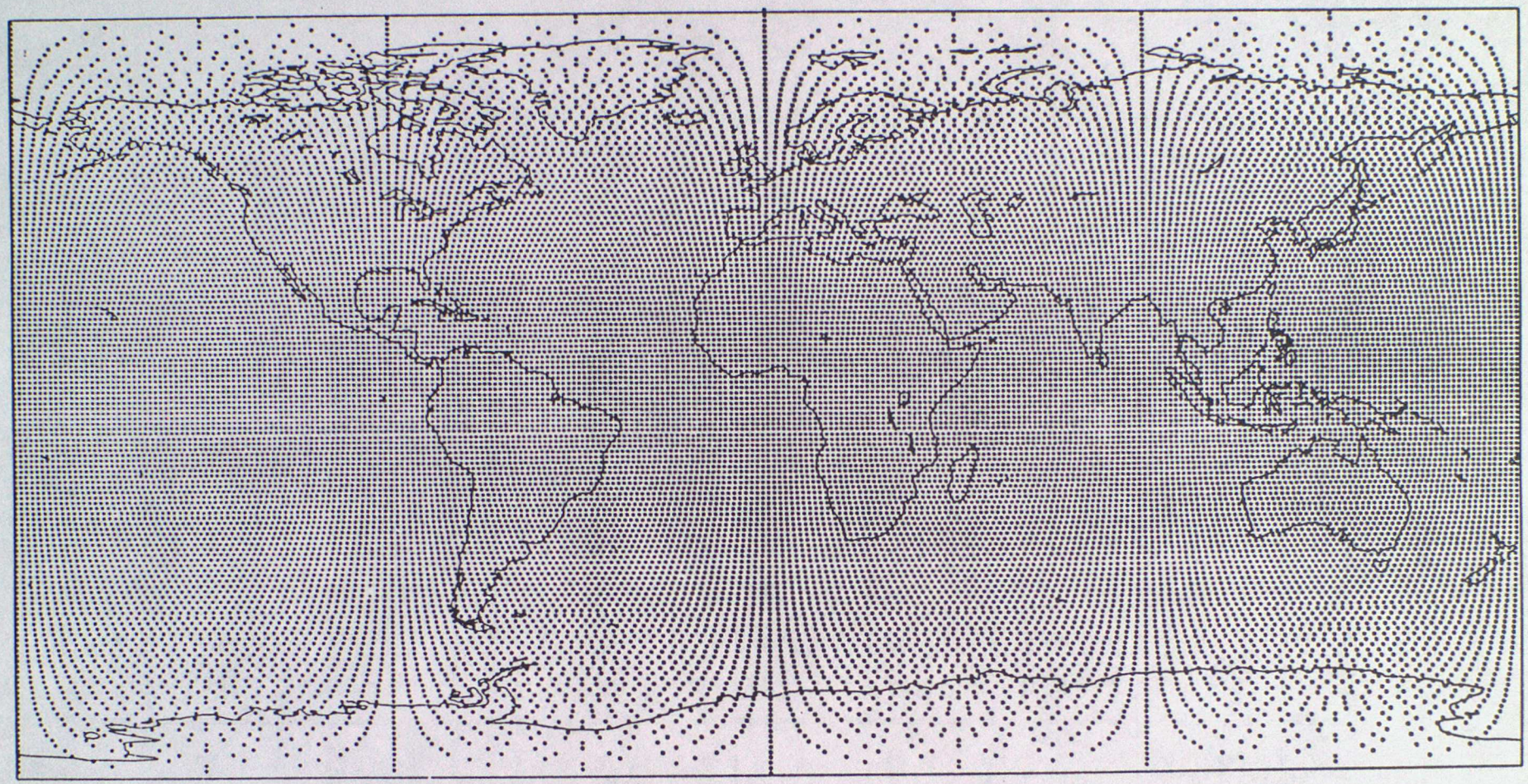
Using second order packing, or row by row, the precision choice above gives the thinned grid 3.29 MBytes, 1.25x1.25 4.82 MBytes and the 1.25x2.5 has 2.52 MBytes. Reducing the precision to 0.5 units, gives an extra saving of 0.44, 0.69 and 0.34 MBytes respectively.

Using the thinned grid wil allow airlines to have a resolution advantage over 1.25x2.5 grid between the 60N and 60S latitudes. This is where the two grids have the same number of points. Above these latitudes, allowing more points does not confer any real advantage, for the distance between points is too short.

It should be remembered that the thinned grid, while giving most of the advantages of the higher resolution (of 1.25x1.25), also has to be re-interpolated to a latitude longitude grid to allow the airlines to do simple interpolation to their routes. Alternatively a more general interpolation routine could be written to allow the airlines to work directly from the thinned grid, which is likely to be less attractive to the airlines because of the complication.

G.H.Ross
P.J.Smith
S.D.Turner
Met O CF6
October 1992

PROPOSED THINNED GRID
FOR WAFS



NUMBER OF BYTES REQUIRED FOR A 1.25 X 2.5 GRID

FOR AN 850 MB LEVEL

NUMBER OF GRID POINTS 20880

FIELD TYPE	ACCURACY	SIMPLE PACKING	MAPPED REPETITION	ROW BY ROW PACKING	SIMPLE PACKING FIXED BIT
Height	1	25826	26250	21480	30974
	0.5	28400	29168	24088	12 bits
	0.25	30974	31576	26680	accuracy=0.25
	0.125	33550	33986	29308	
Temperature	1	18104	12096	12436	20678
	0.5	20678	16942	15008	8 bits
	0.25	23252	21678	17600	accuracy=0.5
	0.125	25826	25824	20192	
Wind U comp	1	18104	14908	14068	20678
	0.5	20678	19648	16640	8 bits
	0.25	23252	23770	19288	accuracy=0.5
	0.125	25826	27218	21880	
Wind V comp	1	15530	13954	13976	20678
	0.5	18104	18010	16512	8 bits
	0.25	20678	21644	19180	accuracy=0.25
	0.125	23252	24876	21768	

NUMBER OF BYTES REQUIRED FOR A 1.25 X 2.5 GRID

FOR A 250 MB LEVEL

NUMBER OF GRID POINTS 20880

FIELD TYPE	ACCURACY	SIMPLE PACKING	MAPPED REPETITION	ROW BY ROW PACKING	SIMPLE PACKING FIXED BIT
Height	1	28400	29382	22568	30974
	0.5	30974	32250	25124	12 bits
	0.25	33550	34716	27788	accuracy=0.5
	0.125	36124	37182	30344	
Temperature	1	15530	8978	10024	20678
	0.5	18104	13536	12344	8 bits
	0.25	20678	18372	15044	accuracy=0.25
	0.125	23252	22774	17620	
Wind U comp	1	18104	16384	15716	20678
	0.5	20678	20472	18272	8 bits
	0.25	23252	24204	20880	accuracy=0.5
	0.125	25826	27518	23492	
Wind V comp	1	18104	17212	16224	20678
	0.5	20678	20962	18832	8 bits
	0.25	23252	24506	21444	accuracy=0.5
	0.125	25826	27640	24052	

Table 1

NUMBER OF BYTES REQUIRED FOR A 1.25 X 1.25 GRID

FOR AN 850 MB LEVEL

NUMBER OF GRID POINTS 41760

FIELD TYPE	ACCURACY	SIMPLE PACKING	MAPPED REPETITION	ROW BY ROW PACKING	SIMPLE PACKING FIXED BIT
Height	1	51566	49274	40636	61862
	0.5	56714	55864	45784	12 bits
	0.25	61862	60466	50968	accuracy=0.25
	0.125	67012	65068	56152	
Temperature	1	36122	17868	22600	41270
	0.5	41270	27326	27620	8 bits
	0.25	46418	37872	32880	accuracy=0.5
	0.125	51566	47786	38060	
Wind U comp	1	36122	22766	26300	41270
	0.5	41270	33472	31408	8 bits
	0.25	46418	43612	36592	accuracy=0.5
	0.125	51566	52294	41796	
Wind V comp	1	30974	22538	25864	41270
	0.5	36122	31700	31012	8 bits
	0.25	41270	40442	36284	accuracy=0.25
	0.125	46418	48026	41380	

NUMBER OF BYTES REQUIRED FOR A 1.25 X 1.25 GRID

FOR A 250 MB LEVEL

NUMBER OF GRID POINTS 41760

FIELD TYPE	ACCURACY	SIMPLE PACKING	MAPPED REPETITION	ROW BY ROW PACKING	SIMPLE PACKING FIXED BIT
Height	1	56714	56026	43900	61862
	0.5	61862	62242	49064	12 bits
	0.25	67021	66992	54264	accuracy=0.5
	0.125	72160	71742	59432	
Temperature	1	30974	12920	18940	41270
	0.5	36122	20472	23688	8 bits
	0.25	41270	30562	28964	accuracy=0.25
	0.125	46418	40954	34076	
Wind U comp	1	36122	27144	29688	41270
	0.5	41270	36610	34852	8 bits
	0.25	46418	45262	40056	accuracy=0.5
	0.125	51566	53188	45204	
Wind V comp	1	36122	29942	30248	41270
	0.5	41270	38618	35380	8 bits
	0.25	46418	46620	40636	accuracy=0.5
	0.125	51566	53840	45784	

Table 2

NUMBER OF BYTES REQUIRED FOR A 1.25 X 1.25 IRREGULAR GRID

FOR AN 850 MB LEVEL

NUMBER OF GRID POINTS 26610

FIELD TYPE	ACCURACY	SIMPLE PACKING	ROW BY ROW PACKING	SIMPLE PACKING FIXED BIT
Height	1	33636	27416	40288
	0.5	36962	30752	12 bits
	0.25	40288	34068	accuracy=0.25
	0.125	43616	37396	
Temperature	1	23658	15868	26984
	0.5	26984	19344	8 bits
	0.25	30310	22604	accuracy=0.5
	0.125	33636	25992	
Wind U comp	1	23658	18340	26984
	0.5	26984	21692	8 bits
	0.25	30310	24956	accuracy=0.5
	0.125	33636	28284	
Wind V comp	1	20332	17948	26984
	0.5	23658	21252	8 bits
	0.25	26984	24564	accuracy=0.25
	0.125	30310	27904	

NUMBER OF BYTES REQUIRED FOR A 1.25 X 1.25 IRREGULAR GRID

FOR A 250 MB LEVEL

NUMBER OF GRID POINTS 26610

FIELD TYPE	ACCURACY	SIMPLE PACKING	ROW BY ROW PACKING	SIMPLE PACKING FIXED BIT
Height	1	36962	28860	40288
	0.5	40288	32148	12 bits
	0.25	43616	35512	accuracy=0.5
	0.125	46942	38800	
Temperature	1	20332	13228	26984
	0.5	23658	16128	8 bits
	0.25	26984	19612	accuracy=0.25
	0.125	30310	22820	
Wind U comp	1	23658	20684	26984
	0.5	26984	23920	8 bits
	0.25	30310	27316	accuracy=0.5
	0.125	33636	30584	
Wind V comp	1	23658	21188	26984
	0.5	26984	24508	8 bits
	0.25	30310	27844	accuracy=0.5
	0.125	33636	31168	

Table 3

Forecast Transmission Sizes

11 Vertical Levels.
3 Parameters (Temp, U, V at 0.25 precision)
4 Forecast times.

	1.25x1.25 Thinned	1.25x1.25	1.25x2.5	2.5x2.5	2.5x5.0
No Pts per globe	26616	41760	20880	10512	5256
Size (in Million Bytes) Based on 250mb test GRIB Packing					
1st Order Packing	3.85	5.90	2.96	1.47	0.74
2nd Order Packing	3.29	4.82	2.52	1.28	0.64
Saving per precision bit	0.44	0.69	0.34	0.17	0.09

This is based on a forecast in the Northern Hemisphere summe. The Annual Fluctuation is less than 10% with the maximum when the Southern Hemisphere jet is strongest.

Different vertical levels have different sizes. Estimates based on 850mb fields are 3% smaller (Temperature Fields larger, Wind smaller). Estimate based on 250mb fields will be slightly too large.

Table 4